(12) UK Patent Application (19) GB (11) 2 274 372 (13) A

(43) Date of A Publication 20.07.1994

(21) Application No 9225212.1

(22) Date of Filing 02.12.1992

(71) Applicant(s)
International Business Machines Corporation

(Incorporated in USA - New York)

Armonk, New York 10504, United States of America

(72) Inventor(s)
Dov Ramm
Dan Chazan

(74) Agent and/or Address for Service
F N Blakemore
IBM UK Ltd, Intellectual Property Dept, Hursley Park, Winchester, Hampshire, SO21 2JN, United Kingdom

(51) INT CL^5
G10K 11/16

(52) UK CL (Edition M)
H4J JGA
U1S S2215

(56) Documents Cited
None

(58) Field of Search
UK CL (Edition L) H4J JGA
INT CL^5 G10K 11/16
Online databases: WPI, Inspec

Best Available Copy

(54) Adaptive noise cancellation device

(57) An adaptive noise cancellation device comprises: convolution logic 10 for convolving the signal from a reference input x with a discretised L-tap filter to form a filtered reference signal; and logic 20 for subtracting the filtered reference signal from a signal input to form an output signal; logic for generating the filter taps as a linear combination of N basis functions each having a corresponding coefficient Ck; and logic for repeatedly determining the coefficients Ck which minimise the power in the output signal z, characterised in that N is less than the number of filter taps L and the basis functions have a portion of finite width, outside of which portion the functions are substantially zero, both in the frequency and time domains, in an embodiment they are gaussians. A full-duplex speakerphone is disclosed including such a noise cancellation device.

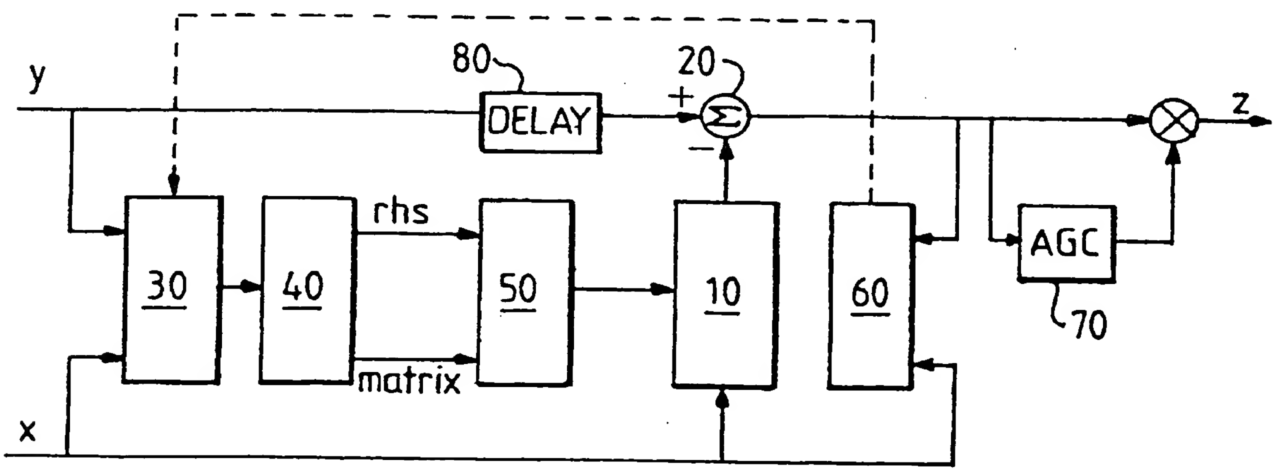

FIG. 1

GB 2 274 372 A

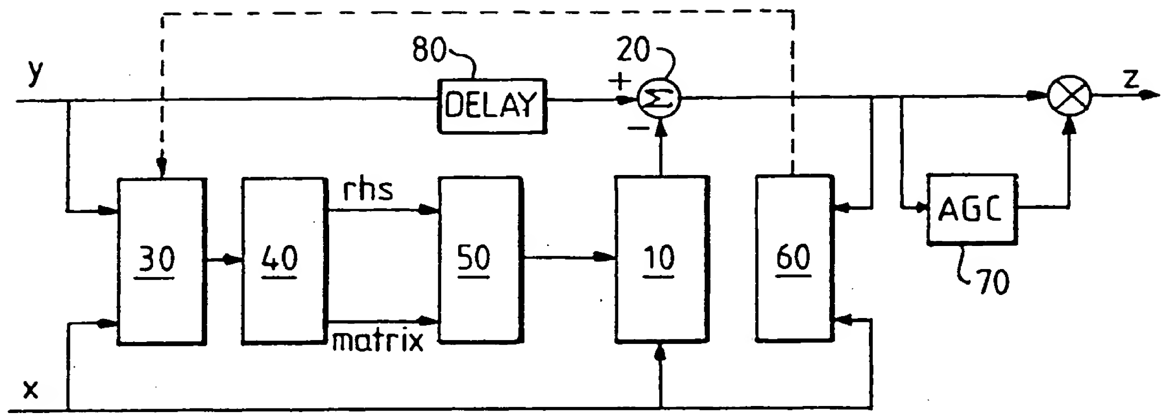
y
80
DELAY
20
+
Σ
−
z
rhs
30
40
50
10
60
AGC
70
matrix
x

FIG. 1

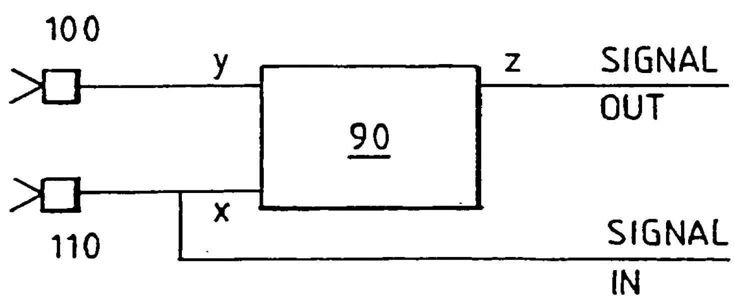
100
y
z
SIGNAL
OUT
90
x
110
SIGNAL
IN

FIG. 2

2274372

ADAPTIVE NOISE CANCELLATION DEVICE

The invention relates to adaptive noise cancellation devices and to the use of adaptive noise concellation devices as howling cancellers for full duplex speakerphones.

It is common practice to use adaptive noise cancellation devices in the transmission of intelligible speech from an environment with considerable acoustic noise, such as an aircraft cockpit, a car etc., utilising as a reference signal the signal from an additional microphone, located sufficiently far from the speaker to be insensitive to the speech of the speaker and as close as possible to the noise source. Adaptive noise cancellation devices are also used in full-duplex speakerphone applications to prevent howling resulting from the feedback path between speaker and microphone of the speakerphone device.

Various techniques for adaptive noise cancelling are known, some of these are described in a paper by B Widrow et al, "Adaptive Noise Cancelling - Principles and Application of the LMS adaptive filter", Proc IEEE, Vol 63, No 12, Dec 1975, pp 1692-1716, and in a paper by H R Sambur, "Adaptive Noise Cancelling for Speech Signals", IEEE trans. ASSP, vol 26, 1977, pp 419-423.

An adaptive noise canceller first produces an estimate of the characteristics of the transformation from the reference noise signal to the noise component of the signal at the main microphone. This transformation actually depends on two acoustic coupling paths: the first between the noise source to the additional microphone, the second between the noise source to the main microphone. Then the reference noise signal is used to model the noisy signal component at the main microphone. This noise component is subtracted from the actual microphone signal. In the absence of the speaker signal and assuming that the transformation is exactly identified, the difference between the actual main microphone input and the estimated noise at the adaptive filter output would be zero. In the presence of the speaker signal this difference signal contains mainly the signal from the speaker.

The filter characteristics are dynamically changed for optimum

removal of the noise signal. The two acoustic paths, which define the transformation between the additional and the main microphone are generally nonstationary. The adaptation rate has to be sufficiently high to adaptively track changes in the impulse responses of these paths due to the movement of people or objects in the room, the movement of the noise source or the change in the noise characteristics.

Adaptive filters usually employ the well established "LMS algorithm", which is known also by the name "stochastic gradients", see, for example, Widrow B, et al "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter", Proc IEEE, August 1976, Vol 64, No 8, PP 1151-1162.

According to the LMS algorithm, the output of the adaptive filter is required to be as small as possible in the sense of least-mean-square error, ie the output power is minimised. The LMS algorithm updates N unknown filter coefficients each input sample and produces an approximately optimal solution in O(N) computations. The LMS algorithm works best when the reference signal is a white noise signal. However, for actual signals which differ from a white noise, and are non-stationary, the convergence of the LMS method is very poor.

The required filter length, ie the number of taps N, is determined by the length of impulse responses of the acoustic leakage paths, that is by the reverberation times of the room. Typical reverberation time of most rooms is less than 400 msec. The adaptive filter length must therefore be about 100 to 200 msec, ie in the range of 1000 to 2000 taps for a common choice of sampling frequency of 8 KHz. Not counting the filtering operation itself, the computational requirements for adaptive LMS filter implementation are therefore at least in the range of 16 to 32 million instructions per second - a very substantial amount of computing power.

Another approach to the problem of adaptive noise cancellation is known as the sub-band acoustic echo canceller (SBAEC), see for example a paper by Andre Gilloire, "Experiments with Sub-Band Acoustic Echo Cancellers for Teleconferencing", Proc of IEEE ICASSP 87, April 1987, pp 2141-2144. These have advantages over the standard LMS method both in adaptation rate and in computational complexity. They do, however, have a residual error due to channel interdependence, which has not been taken

into account in the adaptation scheme, and this limits the number of channels and therefore the computational improvement which may be obtained.

This invention provides an adaptive noise cancellation device having a signal input, a reference input and a signal output, the device comprising: convolution logic for convolving the signal from the reference input with a discretised L-tap filter to form a filtered reference signal; and logic for subtracting the filtered reference signal from the signal input to form the output signal; logic for generating the filter taps as a linear combination of N basis functions each having a corresponding coefficient Ck; and logic for repeatedly determining the coefficients Ck which minimise the power in the output signal, characterised in that N is less than the number of filter taps L and the basis functions have a portion of finite width, outside of which portion the functions are substantially zero, both in the frequency and time domains.

The LMS formulation forms the basis for the approach of this invention. However, unlike the prior art, the filter is represented as a sum of N basis functions, where N is less than the filter length. This is a generalisation of the SBAEC. In the latter, the signal is prefiltered by several band filters. However, for each sub-band a conventional LMS algorithm is applied to adapt a number of filter coefficients in each band, the interdependence of the different bands being neglected. By contrast, in the approach of this invention there are as many "bands" as degrees of freedom. In each sub-band there is only one filter coefficient to be defined and interdependence of adjacent bands is systematically accounted for.

Preferably the basis functions have a small time-bandwidth product. That means that the basis functions have to be well concentrated in the time domain, ie their time-width should be as small as possible and simultaneously these functions have to have a bandwidth in the frequency domain as small as possible.

In practice what this means is that the basis functions are such that within the desired accuracy, both the basis functions and their Fourier transforms may be approximated by narrow functions with finite supports. The use of such basis functions will mean that the filter

which is obtained by combining them, will in general be non-causal, resulting in some additional delay. However, by a judicious choice of parameters this delay can be made equal to only a small portion of the filter length, and small compared to other delays in the system.

In one embodiment, the basis functions are generated by an equation of the following form,

$$w_k(t) = w(t)\exp(jk(\Delta\omega)t) \tag{1}$$

such as are used in a discrete Fourier transform. Here, $j = \sqrt{-1}$, k is an index which runs from -N/2 to (N/2)-1, and $\Delta\omega$ is the step in the frequency domain. In this form the requirement of small time-bandwidth product is applied actually only to the window function w(t).

The requirement that the basis functions have small time bandwidth product results in negligible interdependence between basis functions with indices i and k where $|i-k|$ is larger than about 3. Therefore the system of equations for filter coefficients, which is obtained by a differentiation of the expression for mean-square error, has a sparse matrix, in which only diagonals near the main diagonal differ from zero. Such system can be solved efficiently with O(N) computations. Since a computation of the system matrix itself requires a calculation of correlation functions, which is performed in O(NlogN) operations, the whole filter adaptation can be made in O(NlogN) operations per computational block, yielding a new accurate update of the filter coefficients per block, assuming that the block length is some small integer multiple of the filter length.

An embodiment of the invention will now be described, by way of example only, with reference to the accompanying drawings, wherein

Fig. 1. is a schematic block diagram of the adaptive noise cancellation device of the embodiment of the invention;

Fig. 2. is a schematic block diagram of a speakerphone which includes an adaptive noise cancellation device as a howling canceller.

A block diagram of the adaptive noise cancellation device is shown in Figure 1. The device accepts as input the audio signals reference input signal x and main input signal y, and produces as output the signal z. The device includes the following elements: the convolution logic 10, which convolves the reference signal x with the current filter coefficients; the subtractor 20, which subtracts the convolver output from the main input signal y (appropriately delayed); the correlation logic 30, which computes the autocorrelation function of the reference signal x and the cross-correlation function between the reference and the main input signal; the equation set logic 40, which computes, from these correlation functions, the coefficients of the system of linear equations for filter coefficients; the equation solver 50, which solves the system of linear equations; and the control logic 60, which computes short time power or amplitude estimates for the reference and the subtractor output signals correspondingly and determines a control parameter for the correlation block. Finally, the automatic gain control (AGC) block 70 multiplies the subtractor output by slowly changing gain, the output of the AGC block 70 being the output of the device.

In this embodiment of the invention the logical elements are implemented by a suitably programmed combination of an IBM Personal System/2 personal computer, a plug-in accelerator card which includes an i860 processor chip manufactured by Intel, and an IBM Audio Capture and Playback card for inputting and outputting the signals (IBM and Personal System/2 are trademarks of International Business Machines Corporation). Using this arrangement it has proved possible to attain real-time performance with filters of length 1024.

However, the invention could equally be implemented by a suitably programmed general purpose Digital Signal Processor which performs real time processing of incoming digital signals, or by using special purpose hardware. It will be understood that the blocks indicated schematically in Figure 1 and the above description could in practice be intermixed. For example, the convolution logic could be implemented in the frequency domain and could use the same Fourier transform data and algorithms, as those which are used by the correlation logic.

The reference signal x is fed to the device through a suitable analog/digital (A/D) converter. The input signal y is connected through another A/D converter to the device.

The convolution logic 10 generates a convolution of the reference signal, x, with a filter response function Q which simulates the acoustic path between the reference noise source and the main signal input. The output of the convolution logic 10 simulates the interfering noise signal entering the microphone.

According to the LMS (Least-mean-square) formulation, which forms the basis of the present approach, the output of the adaptive filter is required to be as small as possible in the sense of the least-mean-square error. Since the interfering noise and the desired signal are assumed to be uncorrelated signals, the minimisation problem is to determine Q such that $\{|y-Qx|\}$, ie the output power, is minimised, where y is the primary signal at the main microphone and Qx is the output of the filter Q given the input x to the filter. The filter output $Qx = q \otimes x$, where $\otimes$ denotes a convolution operation and q(t) is a filter impulse response.

The filter Q is represented as a sum of N basis functions, with unknown coefficients, C_k. The filter has the form:

$$q(t) = \sum_{k=-\frac{n}{2}}^{k=\frac{n}{2}-1} C_k \Phi_k(t), \qquad \Phi_k(t) = w(t)\exp(jk(\Delta\omega)t), \tag{2}$$

where $\Delta\omega$ is a frequency step, t is time and w(t) is a window function.

In the frequency domain the Fourier transforms, Φ_k, of the basis functions, $\phi_k(t)$, are shifted copies of the Fourier transform $W(\omega)$ of the window function w(t);

$$\Phi_k(\omega) = W(\omega - k\Delta\omega) = \Phi_0(\omega - k\Delta\omega) \tag{3}$$

Hence, the filter frequency response is found to be a combination of shifted window responses:

$$Q(\omega) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} C_k W(\omega - k\Delta\omega) \qquad (4)$$

The solution of the minimisation problem leads to a set of N linear equations for C_k:

$$\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} a_{ik} C_k = r_i \qquad i = -\frac{N}{2}, \ldots, \frac{N}{2}-1, \qquad (5)$$

with the system coefficients a_{ik}, and the right-hand side r_k. If a_{ik} and r_i can be determined, then C_k can be calculated.

In the general case, the system coefficients a_{ik} can be expressed as follows;

$$a_{ik} = \int_{-\infty}^{\infty} \Phi_I^*(\omega)\Phi_k(\omega)X^*(\omega)X(\omega)\,d\omega \qquad (6)$$

$$= \int_{-\infty}^{\infty} X^*(\omega)X(\omega)W^*(\omega - i\Delta\omega)W(\omega - k\Delta\omega)\,d\omega, \qquad (7)$$

where $X(\omega)$ and $Y(\omega)$ are the Fourier transforms of x and y respectively and the right-hand-side coefficients r_i can be expressed as follows.

$$r_i = \int_{-\infty}^{\infty} \Phi_I^*(\omega)X^*(\omega)Y(\omega)\,d\omega \qquad (8)$$

$$= \int_{-\infty}^{\infty} X^*(\omega)Y(\omega)W^*(\omega - i\Delta\omega)\,d\omega, \qquad (9)$$

By denoting a combination of window responses as:

$$K(\omega,\mu) = W^{*}(\omega-\mu\Delta\omega)W(\omega+\mu\Delta\omega) \qquad (10)$$

and correspondingly an auxiliary function in the time domain:

$$k(t,\mu) = \mathcal{F}^{-1}[K(\omega,\mu)], \qquad (11)$$

the coefficients a_{ik} and r_i can be transformed:

$$a_{ik} = G_{xx}\left(\frac{i+k}{2}\Delta\omega, \frac{i-k}{2}\right) \qquad G_{xx}(\omega,\mu) = \mathcal{F}[R_{xx}(\tau)k(\tau,\mu)] \qquad (12)$$

and

$$r_i = G_{xy}(i\Delta\omega) \qquad G_{xy} = \mathcal{F}[R_{xy}(\tau)w(\tau)], \qquad (13)$$

where R_{xx} is the auto-correlation of the reference signal x and R_{xy} is a function of the cross-correlation between the microphone output y and the reference x.

Thus a_{ik} and r_i and the unknown coefficients, C_k, in the filter response function can be determined from auto-correlations and cross-correlations of the input signals. The filter can then be computed by solving equation (5) which minimises the noise interference in the output, z.

The computation of the filter coefficients involves the correlation logic 30, the equation set logic 40 and the equation solver 50. The correlation block 30 updates current auto- and cross-correlation functions for further analysis, the equation set block 40 uses the updated correlations for setting coefficients of the system of equations, and the equation solver block 50 computes the unknown coefficients.

The window function, w(t), which is used to generate the basis

functions in equation (2) is required to have a small time-bandwidth product. This results in negligible interdependence between basis functions with indices i and k, where $|i-k|$ is larger than about 3, where i and k are complex exponent indices. Thus the system of equations for filter coefficients has a sparse matrix, in which only several, D, diagonals above and below the main diagonal differ from zero. Such a system can be solved with O(N) computations. Since a computation of the system matrix itself requires a calculation of correlation functions, which is performed in O(NlogN) operations, the whole filter adaptation can be made in O(NlogN) operations per computational block, yielding a new accurate update of the filter coefficients per block. Generally the block length should be some small multiple of the filter length, L and the filter length, L, is of the order of a few N, where N is the number of degrees of freedom.

The correlation block subdivides the two signals: the reference x and the main input y, into frames of constant length that are equal to the filter length, L. Odd segments are padded by L zeros on the left and even segments are padded by L zeros on the right.

The Fourier transforms F_{xx}, F_{xy} and F_{yx} of the three one-sided correlations R_{xx}, R_{xy} and R_{yx} all with length L, the lag m being in the range $0 \le m \le L-1$. are initialised to zero. The Fourier transforms of the padded segments are then used to update the Fourier transforms F_{xx}, F_{xy} and F_{yx} as follows;

$$
\begin{aligned}
F_{xx} &= \alpha^2 F_{xx} + \alpha X_{I-1}^* X_I + X_I^* X_I \\
F_{xy} &= \alpha^2 F_{xy} + \alpha X_{I-1}^* Y_I + X_I^* Y_I \\
F_{yx} &= \alpha^2 F_{yx} + \alpha Y_{I-1}^* X_I + Y_I^* X_I
\end{aligned}
\qquad (14)
$$

where X_{I-1} and Y_{I-1} are the Fast Fourier Transforms (FFTs) of the previous segment and X_I and Y_I are the FFTs of the current segment. The "forgetting factor", α, is a real coefficient in the range $0 \le \alpha \le 1$. Its value is determined by the control logic 60.

The output of the correlation logic are three Fourier transforms F_{xx}, F_{xy} and F_{yx}, each containing 2L complex numbers.

The equation set block is activated every few frames. It consists

of a lag domain correlator, a matrix computation block and a right-hand-side computation block.

The lag domain correlator computes the auto-correlation R_{xx} and the cross-correlation R_{xy} in the lag domain $(-(L-1)\leq 0 \leq L-1)$ by using the IFFT (Inverse Fast Fourier Transform) on the outputs F_{xx}, F_{xy} and F_{yx} of the correlation block. The first L numbers out of each 2L are kept. The results are arranged into the auto-correlation, R_{xx}, whose values at negative lags are copied from values of positive lags, since R_{xx} is symmetric and the cross-correlation, R_{xy}, whose values at positive lags are taken as the first L numbers of IFFT of F_{xy} and values at negative lags are taken as the first L numbers of IFFT of F_{yx}.

The cross-correlation, R_{xy}, in the range $-(L-1)$ to $L-1$, are fed into the right-hand-side computation block. The computation of the r_i is based on equation 13 as follows.

The cross-correlation data are multiplied by a sampled window function, and are then padded by zeros to obtain 2L numbers. They are then ordered as follows: $R_{xy}(0), R_{xy}(1), \ldots, R_{xy}(L-1), 0, R_{xy}(-(L-1)), \ldots, R_{xy}(-1)$

The data are transformed to the frequency domain using FFT and the results are decimated by $(L/N):1$, producing N complex numbers r_i.

The auto-correlation, R_{xx}, in the range $-(L-1)$ to $L-1$, are fed into the matrix computation block. The computation of the a_{ij} is based on equation 12 as follows.

The auto-correlation data are multiplied by precomputed samples of one of several auxiliary functions, $k(\tau,\mu)$, which are computed from the window function, for $\mu=0,1/2,1,\ldots,D/2$, where D is the number of diagonals above (and correspondingly below) the main diagonal which differ from zero. The window function $w(t)$ and the frequency step $\Delta\omega$ could be chosen so that D was small.

The D results are padded by a zero in order to obtain 2L numbers for each value of μ. They are then ordered as in the R_{xy} case. The data are transformed to the frequency domain using FFT for each μ.

The FFT results are decimated by $(L/2N):1$, yielding 2N real numbers for each μ. The results for $\mu=0$ are used to fill the main diagonal of the system matrix. The result for $\mu=1/2$ are used to fill the diagonal just above the main one. Complex conjugates of these numbers are used to fill the diagonal just below the main one. The results for the other values of μ are used to fill the next diagonals, until D diagonals above the main one and D diagonals below the main one have been filled.

The resulting sparse system could be solved by using complex arithmetic by Gauss elimination in $O(N.D^2)$ complex operations. In the general case, however, matrix normalisation is performed in the following steps;

1. Matrix symmetrization: The complex system matrix $[a_{ik}]$ is multiplied from the left on the matrix $[a^*_{ki}]$. D additional diagonals will all have small entries of the order of magnitude of those entries which were neglected in the main system. They are thus neglected. So in the resulting real symmetric matrix only 2D+1 diagonals are left.

2. Right-hand-side recomputation: The right-hand-side is multiplied on the matrix $[a^*_{ki}]$.

3. Matrix Regularisation: Theoretically the matrix is always positive definite. However this property can vanish due to truncation errors in computations. Thus it is worthwhile to slightly change the main diagonal entries in order to provide positive definiteness of the system matrix. The maximum number of the main diagonal is multiplied by a small factor, ε, and the result is added to all numbers on the main diagonal.

The above description of the general case has been included to illustrate how the device would work for a general window function. However, it is possible to choose a window function which simplifies the computation. An example of such a function is the gaussian window function,

$$w(t) = \exp\left(-\left((t-t_0)\times\frac{f_s}{\gamma}\right)^2\right) \qquad t_0 = \frac{t_2+t_1}{2} \qquad (15)$$

which the inventors use in the preferred embodiment. This is an

advantageous choice for the window function because it has the minimal possible time-bandwidth product (if that product is defined as a product of the second moments of $w(t)$ and $W(\omega)$). Here f_s is a sampling frequency, γ is a parameter, which is chosen so that the practical support of $w(t)$ is finite; for example, if $\gamma=f_s(t_2-t_1)/6$, then $w(t)\leq 0.0002$ for $t\leq t_1$ and $t\geq t_2$, while $w(t_0)=1$. The centre of the window function is denoted by t_0.

The Fourier transform of this window function is;

$$W(\omega) = \sqrt{\pi}\gamma\exp(-\frac{\gamma^2\omega^2}{4})\exp(\omega t_0) \qquad (16)$$

The Fourier transforms $\Phi_k(\omega)$ of the basis functions $\phi_k(t)$ are shifted gaussians.

When the number, N, of degrees of freedom for the filter increases, the number of gaussians $\Phi(\omega)$ on the frequency axis and their density increase as well. With the correct choice of parameters (γ and $\Delta\omega$) the overlapping between gaussians is small. This means that $\Delta\omega$ must be of the order of $1/N$ and γ must be of the order of N.

The set of parameters used in the embodiment is $\Delta\omega = 2\pi f_s/N$ and γ in the range from 0.25N to 0.75N. The optimal value for γ has been found to be $\gamma = (15/32)N$. However other sets of parameters could be chosen.

With such a choice of parameters, only the diagonals which are near the main diagonal in the matrix $[a_{ik}]$ will have significant values which need to be taken into account in the solution. Typically there are three non-zero diagonals from each side of the main one. The matrix is sparse and can be inverted in O(N) operations.

The general formula for the matrix coefficients a_{ik} yields for a gaussian window function;

$$a_{ik} = C\exp(-\frac{\gamma^2}{8}(\Delta\omega)^2(i-k)^2)G_{xx}(\frac{i+k}{2}\Delta\omega) \qquad G_{xx} = \mathcal{F}[R_{xx}(\tau)\exp(-\frac{1}{2}(\frac{\tau}{\gamma})^2)] \quad (17)$$

and the formula for the right-hand-side coefficients r_j yields for a gaussian window function;

$$r_i = G_{xy}(i\Delta\omega) \qquad G_{xy}(\omega) = \mathcal{F}[R_{xy}(\tau)\exp(-\frac{(t-t_0)^2}{\gamma^2})] \qquad (18)$$

where $C=\gamma\ \pi/2$.

The Fourier transforms are sampled in steps of $\Delta\omega$ in (19) or $\Delta\omega/2$ in (18).

In the case of the gaussian window the steps of matrix symmetrisation and right-hand-side recomputation are unnecessary, and so the calculation of the system matrix is much simpler. Calculation of the matrix $[a_{ik}]$ is based on equation 17 as follows;

The auto-correlation data, R_{xy}, in the range $-(L-1)$ to $L-1$ are multiplied by the sampled gaussian and padded by a zero to obtain 2L numbers. The data are ordered as in the general case. The data are transformed to the frequency domain using FFT. Due to the symmetry of the auto-correlation the result will be real.

The FFT result is decimated by two, yielding 2N real numbers: $F_0, F_1, \ldots, F_{2N-1}$, each of which corresponds to a certain value of i-k and is placed on one of the 2N secondary diagonals of the matrix $[a_{ik}]$. The multiplication by another gaussian, $C\exp\{-\alpha(i-k)^2\}$, $\alpha=(\gamma\Delta\omega)^2/8$, is actually performed along each of the secondary diagonals. For each of the principal diagonals, the value of $|i-k|$ and the value of the gaussian multiplier are constant. The maximum of the gaussian is located on the main diagonal. Due to a practically finite support of the gaussian only a few diagonals are considered as non-zero.

The matrix is defined on a torus, with the right-hand-side connected to the left and the bottom to the top. For a gaussian window the matrix $[a_{ik}]$ is symmetric with purely real entries.

In equations 17 and 18 the FFTs can be used instead of $\mathcal{F}$. The results should be decimated according to the desired step in the frequency domain. To achieve more accurate results a basic resolution of

$\Delta\omega/4$ has been used for the FFTs and consequently a decimation by 4:1 in 17 and 2:1 in 18.

The convolver block utilises the current filter coefficients and computes as its output the convolution of the reference signal, x, with the filter. For this purpose it is necessary to obtain the transversal form of the filter. The previous computation resulted in a set of weights. The transversal form of the filter may be obtained by computing the weighted sum of the basis functions with the recently computed weights. This step may be carried out by using an IFFT followed by multiplication of the result with a sampled window function. Depending on whether the convolver works in the time domain or in the frequency domain it uses the computed impulse response or performs an additional FFT to get the filter frequency response, $Q(\omega)$.

In the preferred embodiment the time domain (the lag domain) is used and the signals (and their correlation functions) are sampled with a period $T=1/f_s$. Hence all expressions in the frequency domain get the form of a discrete Fourier transform (DFT) and the fast Fourier transform (FFT) could be efficiently used for actual computation. This yields the following general expression for the filter taps;

$$q_n = q\left(\frac{n}{f_s}\right) = w\left(\frac{n}{f_s}\right) \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} C_k \exp\left(j\frac{2\pi nk}{N}\right) = Nw_n c_n \qquad \{c_n\} = IFFT\{C_k\} \quad (19)$$

In the preferred embodiment with a gaussian window function, the filter length, L, is chosen to be equal to 2N and a computation of the filter in the time domain is found by putting a zero between every two coefficients C_k and performing a 2N-point IFFT in the equation for the filter taps, yielding the following expression for the filter taps;

$$q_n = q\left(\frac{n}{f_s}\right) = Nc_n \exp\left(-\left(\frac{n}{\gamma}\right)^2\right) \qquad \{c_n\} \quad IFFT\{C_k\} \qquad (20)$$

The delay on the main microphone signal, y, with $t_0=0$ is equal to N taps (half the filter length).

The equation solver block is activated every few frames and the system of N equations with N unknown coefficients C_k is solved. The equation solver uses an appropriate one of the many existing variations of Gauss elimination to solve the system efficiently in O(N) operations.

In the case that the convolver uses the same signal frames as the correlation block, it may use FFTs of the reference x that have already been computed. The convolver can also use smaller signal frames, which, in order to minimise delay, is the case in the preferred embodiment. In that case the convolver should provide the computation of FFTs of smaller signal frames. Also the filter impulse response should be split into several frames and the corresponding FFT calculated.

A filter has to be causal. Therefore in the case that $t_l<0$ the filter is shifted by t_l/f_s taps. In the difference $z=y-Qx$, the main microphone signal, y should therefore be delayed as well.

The device includes a delay 80 for the main microphone signal, which balances any undesirable delay in the adaptive filter due either to incomplete causality of the filter or to delay in the convolver.

The control logic contains either short-time power or amplitude estimators for the reference and the subtractor output. The forgetting factor, α, which is utilised in the correlation block, is set to a certain fixed function of the two control variables; the short-time estimation of the reference, $E^{(x)}$, and the short-time estimation of the subtractor output, $E^{(z)}$.

If the main speaker signal is much larger than the noise component in the main microphone output and $E^{(x)}/E^{(z)}$ is large, α should be set to 1, since past segments of the signal contain information relevant for the current filter adaptation. If the interfering noise signal is dominant and $E^{(x)}/E^{(z)}$ is small, the current signal behaviour bears more relevance on the filter adaptation than past signal behaviour, and so α is set to a value $0\leq\alpha<1$. This also provides a faster adaptation. Correspondingly α could be chosen to be a certain function of the ratio $E^{(x)}/E^{(z)}$. However if both $E^{(x)}$ and $E^{(z)}$ are estimated to be small (i.e. both the microphones are at low noise level) then the set $\alpha=1$ is preferable, since it freezes the filter components.

The adaptive filter is capable of fast adaptation because the adaptation time can be made as small as the filter length.

A conventional analog AGC block, which controls the main microphone output before A/D conversion could be used in the device without additional modifications, since the adaptation rate of the proposed noise cancellation device is sufficient to accommodate slow gain changes. A digital AGC can be used in cases for which an analog AGC at the main microphone is absent. This digital AGC block utilises the average power or amplitude estimate of the subtractor output. Either a separate estimator could be used or the estimator in the control block could be utilised, possibly after additional low-pass (optionally non-linear) filtering. This non-linear filtering could provide for different time constants for rise-time and fall-time. The gain G is defined as a fixed positive function of the output E of the estimator. The larger E is, the smaller the gain G is, and vice versa. For small signals $E \approx 0$ the gain is limited by a certain constant G_0. The subtractor output is then multiplied by the gain to produce the output.

In addition logic can be provided that digitally performs signal compression and decompression; i.e. the incoming signal could be decompressed from a non-linear (usually logarithmic) form to a linear form and the outgoing signal could be compressed from a linear to a non-linear form.

The noise cancellation device can be used as an echo or hawling canceller for a full duplex speakerphone. In this application the speaker phone is arranged as shown in Figure 2 with the signal in from the remote telephone fed to loudspeaker 110 and in addition used as reference signal x. The signal from the microphone 100 is used as input signal y and the adaptive noise cancellation device 90 cancels any feedback signal from the loudspeaker to the microphone.

CLAIMS

1. An adaptive noise cancellation device having a signal input (y), a reference input (x) and a signal output (z), the device comprising:

convolution logic (10) for convolving the signal from the reference input (x) with a discretised L-tap filter to form a filtered reference signal; and

logic (20) for subtracting the filtered reference signal from the signal input to form the output signal;

logic for generating the filter taps as a linear combination of N basis functions each having a corresponding coefficient Ck; and

logic for repeatedly determining the coefficients Ck which minimise the power in the output signal (z),

characterised in that

N is less than the number of filter taps L and the basis functions have a portion of finite width, outside of which portion the functions are substantially zero, both in the frequency and time domains.

2. An adaptive noise cancellation device as claimed in claim 1 wherein the N basis functions have the form:

$$w_k(t)=w(t)\exp\{jk(\delta\omega)t\}$$

for N values of an integer k, where w(t) is a window function having a portion of finite width, outside of which portion the function is substantially zero, both in the frequency and time domains.

3. An adaptive noise cancellation device as claimed in claim 2 wherein the window function w(t) is a gaussian function.

4. An adaptive noise cancellation device as claimed in claim 3 wherein the window function has the form:

$$w(t) = \exp\{-(\frac{tf_s}{\gamma})^2\}$$

with $\Delta\omega = 2\pi f_s/N$ and γ in the range 0.25N to 0.75N.

5. An adaptive noise cancellation device as claimed in any preceding claim wherein the logic for generating the filter response function comprises:

correlation logic (30) for generating the cross-correlation between the input signal and the refernce signal and for generating the auto-correlation of the reference signal;

logic (40) for calculating the right hand side of a system of linear equations for the coefficients Ck using the equation

$$r_i = G_{xy}(i\delta\omega),$$

where $G_{xy}(\omega)$ is the Fourier transform of product of the window function w(t) and the cross-correlation between the input signal and the reference signal;

logic (40) for calculating the coefficients of the system of linear equations using the equation,

$$a_{ik} = G_{xx}(\frac{i+k}{2}\delta\omega, \frac{i-k}{2})$$

where $G_{xx}(\omega,\mu)$ is the Fourier transform of the product of the autocorrelation of the reference signal and an auxiliary function

$$k(t,\mu) = \mathcal{F}^{-1}[W^*(\omega-\mu\delta\omega)W(\omega+\mu\delta\omega)]$$

where $W(\omega)$ if the Fourier transform of the window function w(t); and

equation solving logic (50) for solving the equations $a_{ik}C_k=r_i$ for C_k.

6. An adaptive noise cancellation device as claimed in claim 5 wherein the correlation logic (30) comprises:
logic for segmenting the input and reference signals into frames of a length equal to the filter length L;
logic for generating a Fourier transform of the frames;
logic for combining the Fourier transforms of the frames with a current segment of the auto-correlation function and cross-correlation functions weighted by a factor α, where α is between 0 and 1, to form an updated segment of the cross-correlation and auto-correlation functions.

7. An adaptive noise cancellation device as claimed in claim 6 comprising control logic (60) for calculating the factor α as an increasing function of the ratio of the power or amplitude of the reference signal and the output signal.

8. Adaptive noise cancellation device as claimed in any preceding claim wherein N is half the filter length.

9. A speakerphone capable of full-duplex operation having a signal input and a signal output and comprising an adaptive noise cancellation device (90) as claimed in any preceding claim;
a microphone (100) arranged to provide the input signal (y) to the noise cancellation device;
and a loudspeaker (110) connected to the signal input of the speakerphone, wherein the signal input of the speakerphone provides the reference input (x) for the noise cancellation device and the output of the noise cancellation device (z) provides the signal output of the speakerphone.

Patents Act 1977
Examiner's report to the Comptroller under Section 17 (The Search Report)

Application number

GB 9225212.1

Relevant Technical fields

(i) UK Cl (Edition L) H4J (JGA)

(ii) Int Cl (Edition 5) G10K 11/16

Search Examiner

P J EASTERFIELD

Databases (see over)

(i) UK Patent Office

(ii) ONLINE DATABASES: WPI, INSPEC

Date of Search

1 MARCH 1993

Documents considered relevant following a search in respect of claims 1 TO 9

Category (see over)	Identity of document and relevant passages	Relevant to claim(s)
	NONE	

~~IP - doc99\fil000566~~

SF2(p)

BNSDOCID: <GB___2274372A__I_>

Category	Identity of document and relevant passages	Relevant to claim(s)

Categories of documents

X: Document indicating lack of novelty or of inventive step.

Y: Document indicating lack of inventive step if combined with one or more other documents of the same category.

A: Document indicating technological background and/or state of the art.

P: Document published on or after the declared priority date but before the filing date of the present application.

E: Patent document published on or after, but with priority date earlier than, the filing date of the present application.

&: Member of the same patent family, corresponding document.

Databases: The UK Patent Office database comprises classified collections of GB, EP, WO and US patent specifications as outlined periodically in the Official Journal (Patents). The on-line databases considered for search are also listed periodically in the Official Journal (Patents).

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☑ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.